

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Markussen, Jan
Jonassen, Ib
Havelund, Svend
Brandt, Jakob
Kurtzhals, Peter
Hansen, Hertz Per
- (ii) TITLE OF INVENTION: INSULIN DERIVATIVES
- (iii) NUMBER OF SEQUENCES: 26
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Novo Nordisk of North America, Inc.
 - (B) STREET: 405 Lexington Avenue, 64th Floor
 - (C) CITY: New York
 - (D) STATE: New York
 - (E) COUNTRY: United States of America
 - (F) ZIP: 10174-6401
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: to be assigned
 - (B) FILING DATE: 17-SEPT-1997
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Lambiris, Elias J.
 - (B) REGISTRATION NUMBER: 33,728
 - (C) REFERENCE/DOCKET NUMBER: 4341.204-US
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 212-867-0123
 - (B) TELEFAX: 212-878-9655

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Gly	Ile	Val	Glu	Gln	Cys	Cys	Thr	Ser	Ile	Cys	Ser	Leu	Tyr	Gln	Leu
1				5					10					15	
Glu Asn Tyr Cys Xaa															
20															

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 amino acids
 - (B) TYPE: amino acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

(2) INFORMATION FOR SEQ ID NO:3:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

(2) INFORMATION FOR SEQ ID NO:4:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

(2) INFORMATION FOR SEQ ID NO:5:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

(2) INFORMATION FOR SEQ ID NO:6:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

- 36 -

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 5 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Ser Asp Asp Ala Arg
1 5

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 7 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Thr Lys Ser Asp Asp Ala Arg
1 5

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 6 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Lys Ser Asp Asp Ala Arg
1 5

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 112 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CCAAGTACAA AGCTTCAACC AAGTGGGAAC CGCACAAAGTG TTGGTTAACG AATCTTGTAG 60
CCTTTGGTTC AGCTTCAGCT TCAGCTTCTT CTCTTTTATC CAAAGAAACA CC 112

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 27 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TAAATCTATA ACTACAAAAA ACACATA

27

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 79 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TTGGTTGAAG CTTTGTACTT GGTTCGCGT GAAAGAGGTT TCTTCTACAC TCCTAAGTCT 60

GACGATGCTA GAGGTATTG 79

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TTAATCTTAG TTTCTAGAGC CTGCGGG 27

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 85 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TTGGTTGAAG CTTTGTACTT GGTTCGCGT GAAAGAGGTT TCTTCTACAC TCCTACCAAG 60

TCTGACGATG CTAGAGGTAT TGTCG 85

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 71 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CACTTGGTTG AAGCTTTGTA CTTGGTTTGC GGTGAAAGAG GTTTCTTCTA CACTAAGTCT 60

GACGATGCTA G 71

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 68 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CACTTGGTTG AAGCTTTGTA CTTGGTTTGC GGTGAAAGAG GTTTCTTCTA CAAGTCTGAC 60
GATGCTAG 68

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 66 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CACTTGGTTG AAGCTTTGTA CTTGGTTTGC GGTGAAAGAG GTTTCTTCAA AGTCTGACGA 60
TGCTAG 66

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 594 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
(B) LOCATION: 109..522

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CTTAAATCTA TAACTACAAA AAACACATAC AGGAATTCCA TTCAAGAATA GTTCAAACAA 60
GAAGATTACA AACTATCAAT TTCATACACA ATATAAACGA TTAAAAGA ATG AGA TTT 117
Met Arg Phe
1
CCT TCT ATT TTT ACT GCT GTT TTA TTC GCT GCT TCC TCC GCT TTA GCT 165
Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser Ala Leu Ala
5 10 15
GCT CCA GTC AAC ACT ACC ACT GAA GAT GAA ACG GCT CAA ATT CCA GCT 213
Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala Gln Ile Pro Ala
20 25 30 35
GAA GCT GTC ATC GGT TAC TCT GAT TTA GAA GGT GAT TTC GAT GTT GCT 261
Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe Asp Val Ala
40 45 50
GTT TTG CCA TTT TCC AAC TCC ACC AAT AAC GGT TTA TTG TTT ATC AAT 309
Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu Phe Ile Asn
55 60 65
ACT ACT ATT GCC TCC ATT GCT GCT AAA GAA GAA GGT GTT TCT TTG GAT 357
Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val Ser Leu Asp
70 75 80
AAA AGA TTC GTT AAC CAA CAC TTG TGC GGT TCC CAC TTG GTT GAA GCT 405
Lys Arg Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala
85 90 95

TTG TAC TTG GTT TGC GGT GAA AGA GGT TTC TTC TAC ACT CCT AAG GCT 453
 Leu Tyr Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Ala
 100 105 110 115
 GCT AAG GGT ATT GTC GAA CAA TGC TGT ACC TCC ATC TGC TCC TTG TAC 501
 Ala Lys Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr
 120 125 130
 CAA TTG GAA AAC TAC TGC AAC TAGACGCAGC CCGCAGGCTC TAGAACTAA 552
 Gln Leu Glu Asn Tyr Cys Asn
 135
 GATTAATATA ATTATATAAA AATATTATCT TCTTTTCTTT AT 594

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 138 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser
 1 5 10 15
 Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala Gln
 20 25 30
 Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe
 35 40 45
 Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu
 50 55 60
 Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val
 65 70 75 80
 Ser Leu Asp Lys Arg Phe Val Asn Gln His Leu Cys Gly Ser His Leu
 85 90 95
 Val Glu Ala Leu Tyr Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr
 100 105 110
 Pro Lys Ala Ala Lys Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys
 115 120 125
 Ser Leu Tyr Gln Leu Glu Asn Tyr Cys Asn
 130 135

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Thr Gly Gly Lys

1

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Thr Glu Gly Lys

1

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Gly Asp Thr Lys

1

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 8 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Gly Thr Lys Ser Asp Asp Ala Arg

1

5

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Gly Lys Ser Asp Asp Ala Arg
1 5

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 85 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

TTGGTTGAAG CTTTGTA CTT GGTTCGCGT GAAAGAGGTT TCTTCTACAC TGGTACCAAG 60
TCTGACGATG CTAGAGGTAT TGTCG 85

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 82 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

TTGGTTGAAG CTTTGTA CTT GGTTCGCGT GAAAGAGGTT TCTTCTACAC CGGTAAGTCT 60
GACGATGCTA GAGGTATTGT CG 82